Validating management simulation models and implications for communicating results to stakeholders

Martin A. Pastoors, Jan Jaap Poos, Sarah B. M. Kraak, and Marcel A. M. Machiels

Pastoors, M. A., Poos, J. J., Kraak, S. B. M., and Machiels, M. A. M. 2007. Validating management simulation models and implications for communicating results to stakeholders. – ICES Journal of Marine Science, 64: 818–824.

Simulations of management plans generally aim to demonstrate the robustness of the plans to assumptions about population dynamics and fleet dynamics. Such modelling is characterized by specification of an operating model (OM) representing the underlying truth and a management procedure that mimics the process of acquiring knowledge, formulating management decisions, and implementing those decisions. We employ such a model to evaluate a management plan for North Sea flatfish proposed by the North Sea Regional Advisory Council in May 2005. Focus is on the construction and conditioning of OMs, key requirements for such simulations. We describe the process of setting up and validating OMs along with its effects on the ability to communicate the results to the stakeholders. We conclude that there is tension between the level of detail required by stakeholders and the level of detail that can be provided. In communicating the results of simulations, it is necessary to make very clear how OMs depend on past perceptions of stock dynamics.

Keywords: communication, flatfish, management strategy evaluation, North Sea, operating model, simulation, stakeholders.

Received 30 June 2006; accepted 3 March 2007; advance access publication 21 May 2007.

M. A. Pastoors, J. J. Poos, S. B. M. Kraak and M. A. M. Machiels: Wageningen-IMARES, PO Box 68, 1970 AB IJmuiden, The Netherlands. Correspondence to M. A. Pastoors: tel: +31 2555 64646; fax: +31 2555 64640; e-mail: martin.pastoors@wur.nl

Introduction

Fisheries management plans allow for the development of longer-term strategies and are increasingly being used in the Northeast Atlantic (CEC, 2001, 2006). When fisheries managers or stakeholder organizations (e.g. Regional Advisory Councils) discuss such plans, they often require an evaluation of the possible consequences before the plans are implemented.

A management plan can be evaluated using a simulation model of the fishery system. The operating model (OM) describes the key processes, based on simplifying assumptions about the interactions among the different components. Examples of such models (Kirkwood, 1997; McAllister *et al.*, 1999) may use simplified dynamics of fish stocks and fleets, while attempting to incorporate sufficient aspects of the complex dynamics of real systems (Butterworth and Punt, 1999; Punt *et al.*, 2002). A management procedure (MP; Butterworth, 2007) consists of data collection, stock status evaluation, harvest control rules (HCRs), and implementation. Fleet behaviour has generally been captured in simplistic assumptions (Kraak *et al.*, 2004), and the biological detail is often relatively plentiful (Kell and Bromley, 2004). The uncertainty about the real system and its dynamics plays a key role in evaluating management strategies. This uncertainty reflects a lack of knowledge of processes such as the stock–recruitment relationship, density-dependence, natural variability, and responses of the fleet to the measures imposed.

North Sea plaice (*Pleuronectes platessa*) and sole (*Solea solea*) are taken mainly in a mixed beam trawl fishery with a bycatch of other demersal species. Management of these two species faces many challenges because their spatial distributions and the selectivity characteristics of the gears used to catch them can have substantial side effects on the effectiveness of the measures taken in achieving the intended management objectives. For example, if the quotas set for the two species are not exhausted in synchrony, over-quota catches may be taken (and discarded or landed illegally) of the species with the most restrictive quota, which generally is depleted first. Moreover, the relative catch opportunities for the two species may differ between years. Because the mixed beam trawl fishery for flatfish presents a relatively tractable example, its problems have been well studied in the past (Rijnsdorp and Pastoors, 1995; Kell *et al.*, 1999, 2003, 2004; Pastoors *et al.*, 2000; Kell and Bromley, 2004; Kraak *et al.*, 2004).

The two target species are managed by national shares in the total allowable catch (TAC), days-at-sea restrictions on the fleet, and technical measures. The spawning-stock biomass (SSB) of plaice declined after the early 1990s and has been just above the limit biomass reference point (B_{lim}) since the mid-1990s (ICES, 2006a). The sole stock has fluctuated markedly in response to the appearance of strong year classes. The two most recent year classes are thought to be poor, which could easily reduce sole SSB below its B_{lim} in the near future (ICES, 2006c).

In 2004, the Commission of the European Communities asked the North Sea Regional Advisory Council (NSRAC) for advice "on the implementation of a recovery plan for North Sea plaice and a long-term management plan for sole". NSRAC (2005) issued its advice in July 2005, focusing only on the recovery of the plaice stock to above 230 000 t, the precautionary biomass (B_{pa}) suggested by ICES (2006c). The advice did not deal with long-term management of sole. Because the Netherlands is a major player in these flatfish fisheries, the Dutch Ministry of Agriculture, Nature Conservation and Food Quality (LNV) requested an evaluation of the likely effects of the NSRAC-proposed recovery plan on the sole stock. The details of these evaluations (Poos *et al.*, 2006) are not the focus of this manuscript; instead, we focus on the process. The main questions we pose are: (i) how has the technical evaluation of the recovery plan been initiated and carried through; (ii) how has the uncertainty in various processes been encapsulated in the OM; and (iii) how have the results been communicated and used?

Proposed NSRAC management plan

The stated objective of the management plan has been formulated for plaice only (NSRAC, 2005):

> "a multi-annual management plan should be adopted for plaice in the North Sea with an initial target of reaching an SSB at the B_{pa} level within 3–5 years with a re-evaluation after 3 years and with the long term aim of exceeding B_{pa}. The plan should be implemented as of the 1st of January 2006. The management plan is aimed at reducing pressure on juvenile plaice and would comprise structural effort reductions accompanied by stability in the TAC for plaice. The multi-annual plan should be accompanied by a monitoring and evaluation scheme, which would also include the monitoring of social and economic impact".

The basic management measure proposed was "a structural effort reduction of 15% of enforced licensed capacity limits in the international 80 mm flatfish fishery over 2006 and effort to be maintained at the new level for a further two years". There were detailed comments in the plan on exemptions for some fleets, on the way effort reductions could be embedded in national fishing plans, and on the different forms of effort reduction possible (days-at-sea regulations or decommissioning).

Despite these technical details, evaluation with the available scientific tools was difficult. For example, the plan stated: "In the event of the plaice stock falling below B_{lim} new measures would be applied". However, the nature of these measures was not described. Furthermore, the state of the sole stock was not considered, and the clients wanted to know about the effects of the plan on sole.

Such open ends appear to be a generic property of the outcome of political negotiation: specific conditions are raised, but how they should be resolved is left open until problems arise. However, for technical evaluation, open ends present difficulty in interpreting a management plan: a simulation approach requires that all actions be specified under all conditions.

When management plans are developed in conjunction with an evaluation approach, the problems can be resolved through a system of feedback loops between analysts and stakeholders. However, when an evaluation of a specific plan is requested, the open ends have to be interpreted by the analysts by formulating a range of potential scenarios.

This happened in this case: we interpreted the clause about "the new measures" in such a way that, if the perceived SSB would fall below B_{lim}, a further decrease of nominal fishing effort by 15% annually would be implemented until the perceived SSB had returned to above B_{lim}. Further, our interpretation was that the rule (although this had not been specified explicitly) would also apply to sole. An alternative scenario evaluated was without additional measures being taken when SSB would fall below B_{lim}.

The evaluation process

The parties involved in the evaluation process were the Dutch Ministry (LNV), the stakeholder flatfish Working Group (NSRAC), and the research organization (Wageningen-IMARES). The approach was to develop a FLR (www.flr-project.org; Kell *et al.*, 2007) simulation model that would allow incorporation of different hypotheses on population dynamics and the joint exploitation of the two stocks. The three parties initially interacted to define the research questions and the type of results expected, and again at the end of the process to discuss the results in the report. Between start and finish, interaction between researchers and stakeholders was negligible.

Setting up the model

The initial question in model development referred to the definition of the dimensions to be included. The dimensions reflect the type of processes included, the available information, and the dominant issues in the management plan. In this case, the partial spatial overlap between the two species suggested a two-area model (sole dominating as the target south of 56°N, plaice dominating north of 56°N), and a distinction between two main types of fleets (one each targeting sole and plaice).

The simulation model consisted of two submodels: an OM describing the biological and fleet (including economics) dynamics of the underlying system; and an MP consisting of data collection, stock assessment, and a HCR (for a graphical presentation of a similar type of simulation approach, see Kell *et al.*, 2005).

The biological processes in the OM included a stock–recruitment relationship, fixed values for natural mortality, maturity-at-age, and weight- and length-at-age, and fixed values for the relative proportion of each age group in each of the two areas. The fleet-dynamics model consisted of two beam trawl fleets: the Dutch fleet targeting primarily sole (80 mm mesh in the codend) south of 55°N, and a UK fleet targeting primarily plaice (100 mm mesh) north of 55°N. Trends in fishing effort were derived from information in the Dutch logbook database (containing records of Dutch vessels as well as UK vessels landing their catches in the Netherlands). Future effort allocations were assumed to retain the same proportions by area and fleet. Conceptually, the catching process was modelled as a combination of catchability, selectivity, effort, and technological creep (Rijnsdorp *et al.*, 2006). Catches were generated from the underlying ("true") population and split into landings and discards according to a retention ogive derived from observer trips (Van Keeken *et al.*, 2004).

The MP consisted of three main processes: sampling raw data from the underlying population; stock assessment, and short-term forecasting following standard procedures; and a HCR defining the appropriate management measure given the forecast. Sampling from the true population was mimicked by generating estimates of landings-at-age (sole) and catch-at-age (plaice; including discards), similar to the annual assessments for the two stocks. The catches were generated using the selectivity characteristics of the two fleets and a simple lognormal error with a relatively small coefficient of variation ($CV = 0.1$).

The simulation contained a "true" survey that sampled from the populations of the two species using catchability and selectivity patterns estimated from historical data in conjunction with their spatial distribution. "Observed" survey catch-at-age by species was generated by applying a lognormal error (again $CV = 0.1$), and these series were used for tuning in the stock assessment process. The stock assessment process encompassed single-species extended survivors analysis (XSA) for plaice and sole, based on catch- and landings-at-age data, respectively. XSA settings and short-term forecasts corresponded to those used by ICES (2006a).

The HCR implemented in the model attempted to mimic the NSRAC management plan. Nominal fishing effort was reduced by 15% in 2006 compared with 2005, and this level was maintained in subsequent years. In the objective stated in the plan, there is an inherent tension between reducing effort and maintaining stability of the TAC. This posed additional challenges to implementation of the model. LNV requested an additional maximum annual change in TAC of 15% to be included in the simulations, representing an extension of the stated objectives. However, the measures stated did not refer to TACs, but only to decommissioning and days-at-sea limits. Therefore, TACs would not constrain the fishery in the model anyway: the fleets simply exhausted the effort quota and reported whatever catches they generated. We did not implement a two-tier system in which either the TAC or the effort quota could constrain the fishery, because the proposed HCR did not specify how the priority between such different measures would have been set.

The annual decision process on effort quota was based on the short-term forecast of the SSB remaining after the year to which these would apply. This forecast was compared with the $B_{\rm lim}$ triggers defined in the plan. Implementation error with respect to misreporting or black landings was not included.

Parameter estimates

The average distribution of the species by age over the two areas was estimated from the annual beam trawl survey, which takes a synoptic sample of the plaice and sole populations in late summer (ICES, 2006b). In general, sole are mainly in the southern area, and plaice migrate from south to north with increasing age.

Information on stock trends was available for the period 1957–2004 from ICES (2006a). Recruitment estimates for the last four years were excluded because they were considered unreliable. The remaining set was used to estimate stock–recruitment relationships of the Ricker and the Beverton–Holt type. In the forward simulation, recruitment estimates were taken from the stock–recruitment relationship, taking into account the variance estimate derived from the historical relationship.

Estimating the parameters of the fleet-dynamics model was less straightforward. The interplay of simulated stock abundance and stock distributions by age and area, together with the distributions of the fleets, determines the catch profiles generated by the fleets. The empirical observations on the catch profiles of the real fleets were not sufficient to determine the parameter values for the simulation model uniquely. Therefore, any mismatch between the simulated and the real catch profiles could have been caused by incorrect assumptions in the fleet-dynamics or stock-dynamics models.

The critical question in interpreting simulation results is: how close do these have to be to reality as currently perceived? In other words, is it a problem if simulated catch profiles or stock trends differ from the historical observations? The answer depends on the type of evaluation process in which the simulation model is used. The management plan has been phrased in a stock assessment type of discourse by referring to SSB and biological reference points. Therefore, the numerical values of the variables derived from the simulations would be viewed critically against the values produced by routine stock assessment (in this case as presented by ICES, 2006a).

The combination of over-parameterization of the OM (using more parameters than there is information to estimate them from) and the requirement to generate realistic stock and fleet dynamics, implied that fixes had to be sought to generate largely comparable results. We used the technique of hindcasting to simulate the historical trajectories of stocks and fleets over the period 1995–2004. The start populations in 1995 and levels of recruitment were taken directly from ICES (2006a), and the trends in fishing effort were derived from the logbook database. Given these fixed inputs, simulations were run under different assumptions on the relative distribution of fish over the two areas, the catchabilities of the two fleets for the two species, and the increase in technical efficiency of the two fleets (Rijnsdorp *et al.*, 2006).

Some results of the exploratory analyses are shown in Figure 1; we compare the perceived values (ICES, 2006a) of landings, discards, recruitment, average fishing mortality (*F*bar), selection pattern, and SSB, the assumed relative distribution of fish over areas, and the relationship between *F*bar and effort with their simulated ("true") values, as derived from the OM. Obviously, if we wish to fix the simulated historical dynamics to match the observations, we are faced with the choice between many different options: mean selection pattern, effort–mortality relationship, landings, etc., may all be fixed, but fixing all simultaneously is impossible unless weights can be assigned to each specific parameter.

Figure 2 (left panels) shows the comparison between the (perceived as) "true" and simulated values from the basic run for four important population dynamics parameters. In the base run, all parameters were based on information from some type of analysis except for overall catchability of the fleets, which was obtained by minimizing the difference in catch-at-age between the model and as actually observed in samples from the fleet. Clearly, the simulated landings and SSB of plaice and sole are substantially smaller than those according to the assessment. Although the temporal dynamics may appear reasonably similar, a discrepancy in the overall level may have substantial effects on prospective simulations if the same absolute values are used for limit reference points. For instance, the simulated stocks will be below $B_{\rm lim}$ most of the time.

For plaice, an additional problem arose with the split of catches between landings and discards: simulated landings were below and simulated discards above the assessment estimates. Most plaice aged 1 and 2 years are in the southern area where the beam trawl fishery mainly operates, resulting in high simulated discards. In practice, these smaller fish may not be available to the fishery because they live in shallow coastal regions where larger vessels are not permitted. However, in a simulation environment based on two areas only, such subtle differences cannot be accounted for.

The results presented in Figure 2 (right panels) incorporate an attempt to remedy this shortcoming by manually modifying the relative spatial distribution of plaice aged 1 and 2. Because these age groups are now located more in the northern area, the overall discards decrease, so the simulated landings approach the observed landings. Also, the SSB estimates are closer. However, the simulations share a major discrepancy between simulations and assessment (Figure 1f): the linear and positive relationship

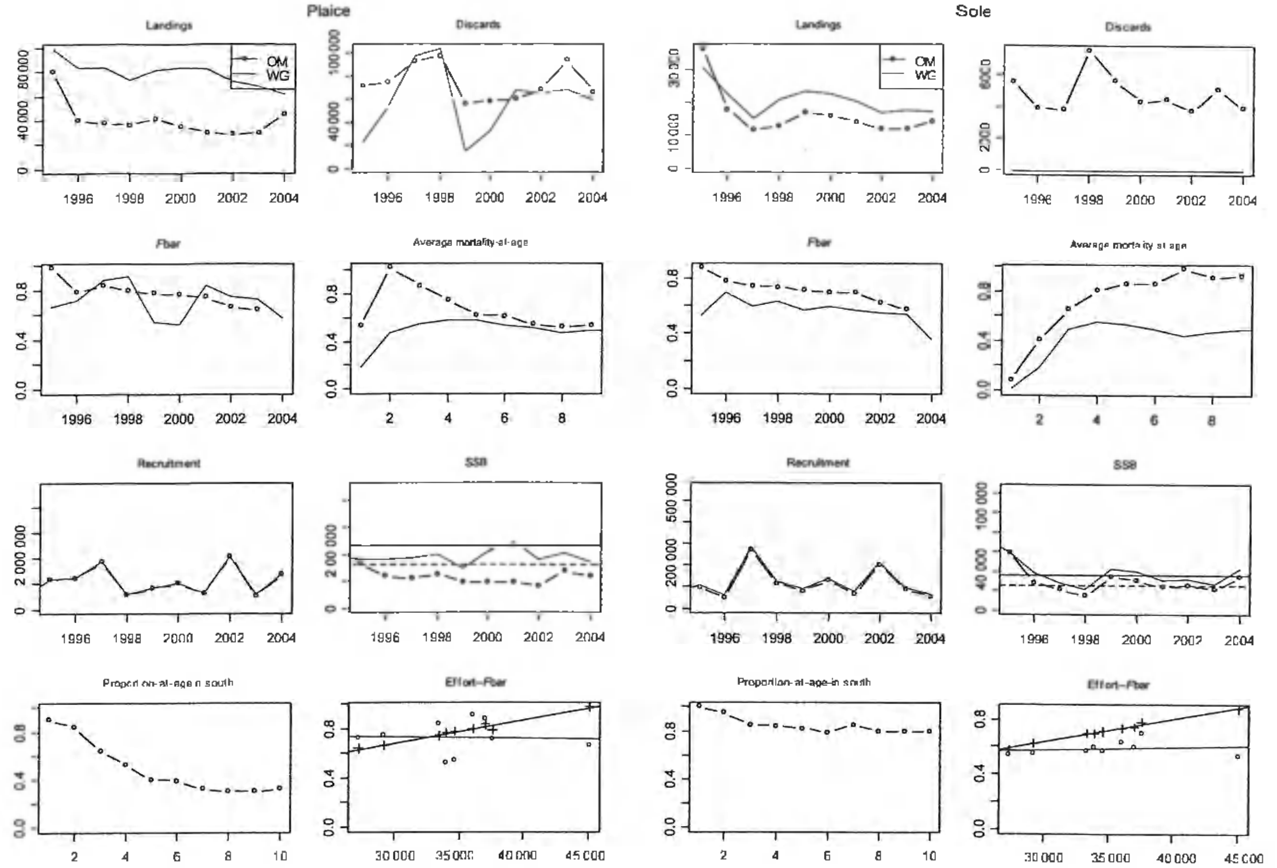

Figure 1. Comparison of hindcast population dynamics parameters for plaice (left) and sole (right) from simulations (black lines, OM) using species distribution by area from surveys and not taking into account technological creep with the corresponding estimates (red lines, Working Group) available from ICES (2006a): landings, discards, mean *F* (Fbar), mortality-at-age, recruitment, SSB, proportion of population by age group in the southern area (input), and the relationship between effort and *F*.

between fishing effort and fishing mortality in the simulation model is a direct consequence of an assumption in the underlying model, whereas the available assessment data indicate no significant relationship at all.

Figure 3 compares the results for sole after introducing an increase in technical efficiency for the two species, as estimated by Rijnsdorp *et al.* (2006). Although the slopes of the effort–mortality relationship have become more comparable, the absolute estimates of fishing mortality in the simulation are still substantially higher than those from the assessment.

Communicating results

Because our focus during the simulations was on the assumptions behind the simulation model and the way the OM was parameterized or, more generally, on the uncertainties involved, the full range of results was presented to the parties that commissioned the study. However, those clients did not relate to the technical complexities of the modelling process. Their main frame of evaluation was whether the hindcasting process made sense (are the results similar to what has been observed?) and whether they could understand the logic of future developments (can the results be explained?). Many stakeholders read the outputs closely: if we simulated landings of plaice of 70 000 t while the TAC was just 60 000 t, they would question the validity of the results. This indicated a large degree of trust in the potential precision of the simulation models used, but at the same time a high degree of scepticism regarding their outcomes: the results were easily considered to be falsified when compared with possibly equally uncertain assessment data.

Discussion

The evaluation of the NSRAC management plan required the use of a multifleet, multistock simulation model, which is a level of complexity higher than most models used so far to evaluate fisheries management systems (Kell *et al.*, 2005, 2006b). The difficulties encountered in trying to match the results of this exercise with the type of information underlying the management advice have not been presented to highlight the modelling details, but to illustrate the general problems of such an evaluation process, of encapsulating uncertainty, and of communicating complex results.

The evaluation was carried through in an almost linear chain of actions. First, the problem was defined jointly by the clients and scientists. During the next phase, the scientists interpreted the plan, developed the model, ran the scenarios, and evaluated the results, almost in isolation. Finally, the results and conclusions were presented to the clients. The lack of interaction with the clients during the actual research phase may be explained partly by the timing of the request, because the work had to be done in December and January when the clients were heavily involved in the annual decision-making process for the following year's

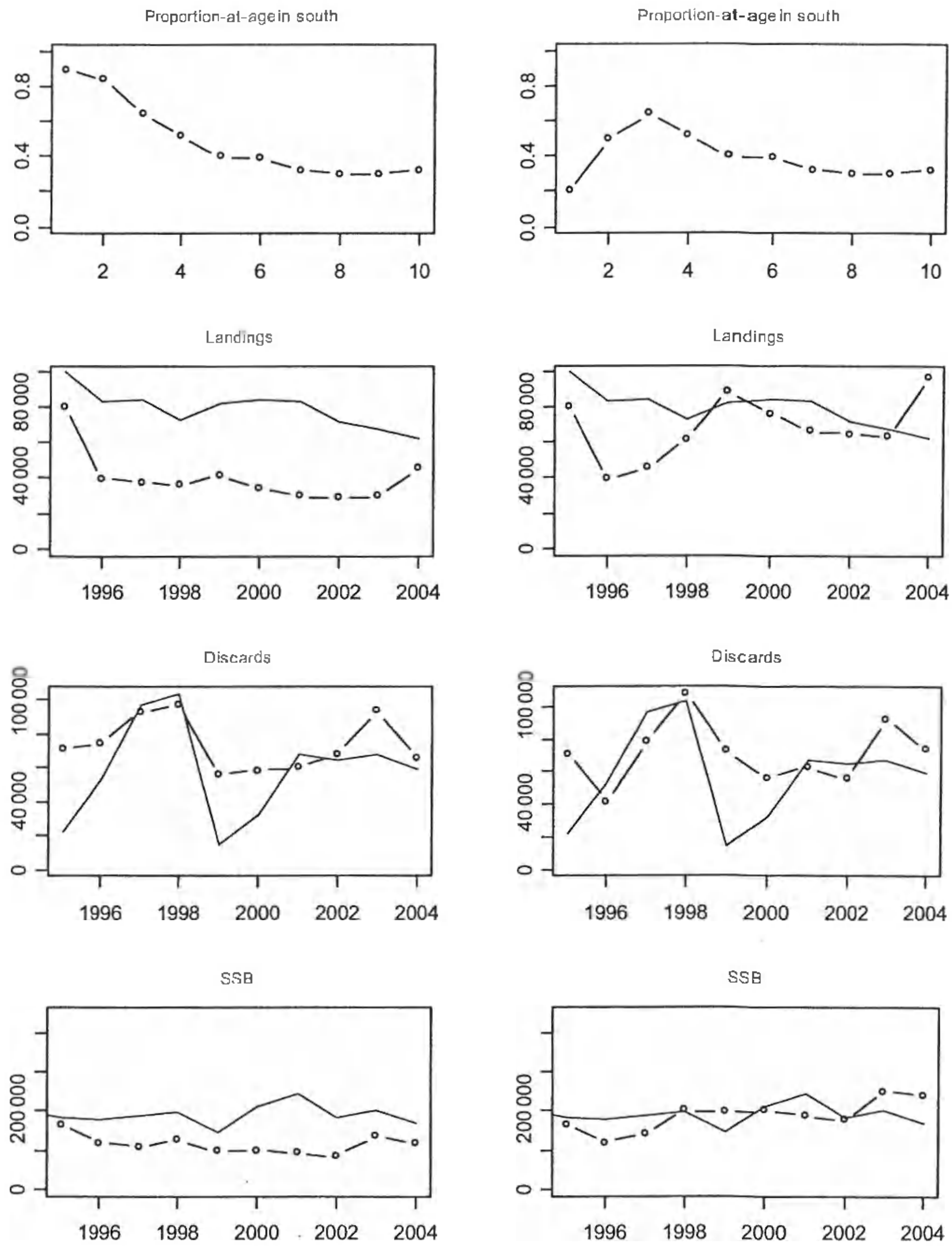

Figure 2. Comparison of hindcast landings, discards, and SSB of plaice for two patterns of the proportion-at-age present in the southern area (top panels): left panels, proportion estimated from survey data (as in Figure 1); right panels, proportion of younger ages manually modified to minimize the discrepancy in the landings.

TAC. Nevertheless, such a lack of interaction made the process less transparent (who takes responsibility for the choices made?), and no adjustments could be made to the mutual expectations of the type of results generated from the simulation model.

A generic property of management plans is that they are the product of political negotiation. Such processes often lead to an end product that describes the overall intention clearly and contains some aspects described in meticulous detail. However, other aspects are only dealt with cursorily, so requiring further interpretation. Also, specific issues may not be tractable in a simulation approach. An example is the apparent willingness of the NSRAC to consider an exception clause for the German shrimper fleet: "if scientific evidence supports the claim that plaice are not discarded in significant quantities in their targeted 80 mm sole fishery". As scientists, we did not see how this exemption could be integrated and evaluated in the simulations without having access to specific information regarding discards in that fishery.

Regarding open ends in a management plan (such as the statement that new measures might be applied under specific conditions, without actually stating what these measures might be), one must accept that the negotiation process cannot simply be re-opened. However, to obtain a clearer understanding of the type of measures envisioned, it would have been beneficial for the

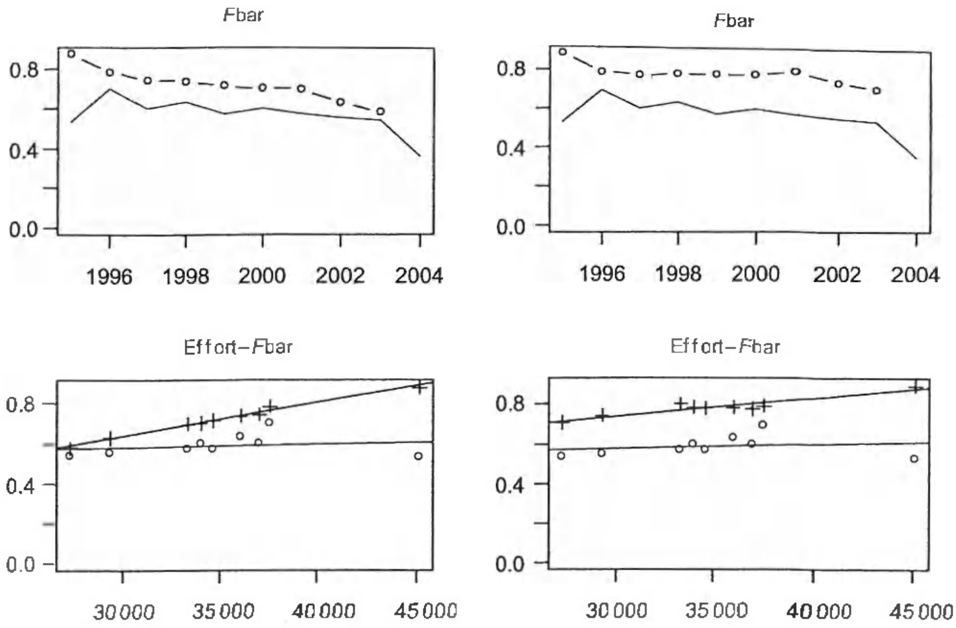

Figure 3. Comparison of hindcast average F (Fbar) and the effort–mortality relationship for sole: left panels, proportion in the south estimated by surveys and no technological creep assumed (as in Figure 1); right panels, proportion in the south manually modified (as for Figure 2) and technological creep assumed as based on Rijnsdorp *et al.* (2006).

evaluation process to have had closer and more regular interactions with the NSRAC on the class of new measures that might be explored.

The uncertainties in the OM explored in simulations addressed just three processes, but these represent a limited subset of all the processes that might influence the overall uncertainty (e.g. selectivity of the fleets, dynamic effort allocations by area, alternative relationships between selectivity, catchability, effort, and fishing mortality). This type of multidimensional model requires a strong analytical approach to explore the main factors affecting uncertainty (Kell *et al.*, 2006a). Hindcasting can be a useful tool for exploring the behavioural dynamics of a model if the dynamics can be verified by observations. As a caveat, we note that the observations available may themselves be flawed interpretations of the reality they attempt to represent. The discard observations in the North Sea plaice assessment are a good example, because the estimates are based on a modelling approach that generates huge fluctuations in discards from year to year. If simulated discards deviate from assessment discards, this may represent incorrect assumptions in either model.

Communication of results with the clients has been addressed only briefly so far. Stakeholders and fisheries managers did not appear to relate directly to the technical complexities of the modelling process, although they were concerned whether the hindcasting could reproduce the observed dynamics in the assessment underlying the TAC advice, and whether they could logically explain in their own words the developments in stocks and fisheries predicted by the simulations. We suggest that the issue of focusing too much on absolute values in the simulation output might be circumvented by expressing the results relative to an appropriately chosen reference level (for instance, the final year for which information is available). The reference points used in the management process would then also have to be scaled on that basis. Even under such a scenario, shared understanding of what the models can and cannot deliver remains an important issue. The expectations of the different parties should be laid clearly and openly on the table, to make the process as efficient as possible.

Our overall conclusion is that the simulation approach has been useful, but it might have been even more useful if used during the developmental stage of the plan, and with more interaction with the NSRAC before the plan was finally agreed.

The FLR modelling environment (Kell *et al.*, 2007) allows construction of flexible simulation models that reuse generic tools and can be attuned to different types of management rules and output measures required. However, because FLR is still in a developmental stage, the balance in the development of generic and specific tools has tended to be on the latter side. As a consequence, the resulting code has become too complex to be transferred easily to other applications. More work is needed to make the simulation models more transparent and transferable.

Acknowledgements

The study was commissioned and funded by the Dutch Ministry of Agriculture, Nature Conservation and Food Quality (LNV). The Commission of the European Communities Directorate General for Fisheries contributed to the development of the simulation model and the writing of the manuscript under the contract "Operational evaluation tools for fisheries management options" ("EFIMAS"; *SSP8-CT-2003-502516*).

References

Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. ICES Journal of Marine Science, 64: 613–617.

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

CEC. 2001. Proposal for a council regulation establishing measures for the recovery of cod and hake stocks. COM (2001), 724.

CEC. 2006. Communication from the Commission to the Council and the European Parliament. Implementing sustainability in EU fisheries through maximum sustainable yield. COM(2006), 360.

ICES. 2006a. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak, Copenhagen, 6–15 September 2005. ICES Document CM 2006/ACFM: 09.

ICES. 2006b. Report of the Working Group on Beam Trawl Surveys, Hamburg, 16–19 May 2006. ICES Document CM 2006/LRC: 11.

ICES. 2006c. Report of the ICES Advisory Committee on Fishery Management, Advisory Committee on the Marine Environment and Advisory Committee on Ecosystems, 2006. ICES Advice, 1–10.

Kell, L. T., and Bromley, P. J. 2004. Implications for current management advice for North Sea plaice (*Pleuronectes platessa* L.). 2. Increased biological realism in recruitment, growth, density-dependent sexual maturation and the impact of sexual dimorphism and fishery discards. Journal of Sea Research, 51: 301–312.

Kell, L. T., Cotter, A. J. R., Van Keeken, O., Pastoors, M., O'Brien, C. M., Piet, G. J., and Rackham, B. D. 2003. The influence on stock assessment advice of sampling error in research survey and international market sampling data for North Sea cod (*Gadus morhua* L.) and plaice (*Pleuronectes platessa* L.). ICES Document CM 2003/X: 5.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E. *et al.* 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640–646.

Kell, L. T., O'Brien, C. M., Smith, M. T., Stokes, T. K., and Rackham, B. D. 1999. An evaluation of management procedures for implementing a precautionary approach in the ICES context for North Sea plaice (*Pleuronectes platessa* L.). ICES Journal of Marine Science, 56: 834–845.

Kell, L. T., Pastoors, M. A., Scott, R. D., Smith, M. T., Van Beek, F. A., O'Brien, C. M., and Pilling, G. M. 2005. Evaluation of multiple management objectives for Northeast Atlantic flatfish stocks: sustainability vs. stability of yield. ICES Journal of Marine Science, 62: 1104–1117.

Kell, L. T., Pilling, G. M., Kirkwood, G. P., Pastoors, M. A., Mesnil, B., Korsbrekke, K., Abaunza, P. *et al.* 2006b. An evaluation of multi-annual management strategies for ICES roundfish stocks. ICES Journal of Marine Science, 63: 12–24.

Kell, L. T., Pilling, G. M., and Kuikka, S. 2006a. Relative importance of value of information and value of control for mixed fisheries. Poster presented at the ICES Symposium on Fisheries Management Strategies, 27–30 June 2006, Galway, Ireland.

Kell, L. T., Scott, R. D., and Hunter, E. 2004. Implications for current management advice for North Sea plaice. 1. Migration between the North Sea and English Channel. Journal of Sea Research, 51: 287–299.

Kirkwood, G. P. 1997. The revised management procedure of the International Whaling Commission. *In* Global Trends: Fisheries Management, pp. 91–99. Ed. by E. K. Pikitch, D. D Huppert, and M. P. Sissenwine. American Fisheries Society Symposium, 20.

Kraak, S. B. M., Buisman, F. C., Dickey-Collas, M., Poos, J. J., Pastoors, M. A., Smit, J. G. P., Daan, N. *et al.* 2004. How can we manage mixed fisheries? A simulation study of the effect of management choices on the sustainability and economic performance of a mixed fishery. ICES Document CM 2004/FF: 11.

McAllister, M. K., Starr, P. J., Restrepo, V. R., and Kirkwood, G. P. 1999. Formulating quantitative methods to evaluate fishery-management systems: what fishery processes should be modelled and what trade-offs should be made? ICES Journal of Marine Science, 56: 900–916.

NSRAC. 2005. Immediate Management Measures for North Sea plaice. Advice to the European Commission. http://www.nsrac.org/advices/wf20050712_Advice_on_plaice_management.pdf.

Pastoors, M. A., Rijnsdorp, A. D., and Van Beek, F. A. 2000. Effects of a partially closed area in the North Sea ("plaice box") on stock development of plaice. ICES Journal of Marine Science, 57: 1014–1022.

Poos, J. J., Machiels, M. A. M., and Pastoors, M. A. 2006. Investigation of some Management Scenarios for North Sea Sole and Plaice in 2006 and beyond. Wageningen-IMARES: CVO Report, 06.004. IJmuiden, the Netherlands.

Punt, A. E., Smith, A. D. M., and Cui, G. 2002. Evaluation of management tools for Australia's South East Fishery 1. Modelling the southeast fishery taking account of technical interactions. Marine and Freshwater Research, 53: 615–629.

Rijnsdorp, A. D., Daan, N., and Dekker, W. 2006. Partial fishing mortality per fishing trip: a useful indicator of effective fishing effort in mixed demersal fisheries. ICES Journal of Marine Science, 63: 556–566.

Rijnsdorp, A. D., and Pastoors, M. A. 1995. Modelling the spatial dynamics and fisheries of North Sea plaice (*Pleuronectes platessa* L.) based on tagging data. ICES Journal of Marine Science, 52: 963–980.

Van Keeken, O. A., Quirijns, F. J., and Pastoors, M. A. 2004. Analysis of Discarding in the Dutch Beam Trawl Fleet. RIVO Report, C034/04. IJmuiden, the Netherlands.

doi:10.1093/icesjms/fsm051

Validating management simulation models and implications for communicating results to stakeholders

Martin A. Pastoors, Jan Jaap Poos, Sarah B. M. Kraak, and Marcel A. M. Machiels

Pastoors, M. A., Poos, J. J., Kraak, S. B. M., and Machiels, M. A. M. 2007. Validating management simulation models and implications for communicating results to stakeholders. – ICES Journal of Marine Science, 64: 818–824.

Simulations of management plans generally aim to demonstrate the robustness of the plans to assumptions about population dynamics and fleet dynamics. Such modelling is characterized by specification of an operating model (OM) representing the underlying truth and a management procedure that mimics the process of acquiring knowledge, formulating management decisions, and implementing those decisions. We employ such a model to evaluate a management plan for North Sea flatfish proposed by the North Sea Regional Advisory Council in May 2005. Focus is on the construction and conditioning of OMs, key requirements for such simulations. We describe the process of setting up and validating OMs along with its effects on the ability to communicate the results to the stakeholders. We conclude that there is tension between the level of detail required by stakeholders and the level of detail that can be provided. In communicating the results of simulations, it is necessary to make very clear how OMs depend on past perceptions of stock dynamics.

Keywords: communication, flatfish, management strategy evaluation, North Sea, operating model, simulation, stakeholders.

Received 30 June 2006; accepted 3 March 2007; advance access publication 21 May 2007.

M. A. Pastoors, J. J. Poos, S. B. M. Kraak and M. A. M. Machiels: Wageningen-IMARES, PO Box 68, 1970 AB IJmuiden, The Netherlands. Correspondence to M. A. Pastoors: tel: +31 2555 64646; fax: +31 2555 64640; e-mail: martin.pastoors@wur.nl

Introduction

Fisheries management plans allow for the development of longer-term strategies and are increasingly being used in the Northeast Atlantic (CEC, 2001, 2006). When fisheries managers or stakeholder organizations (e.g. Regional Advisory Councils) discuss such plans, they often require an evaluation of the possible consequences before the plans are implemented.

A management plan can be evaluated using a simulation model of the fishery system. The operating model (OM) describes the key processes, based on simplifying assumptions about the interactions among the different components. Examples of such models (Kirkwood, 1997; McAllister *et al.*, 1999) may use simplified dynamics of fish stocks and fleets, while attempting to incorporate sufficient aspects of the complex dynamics of real systems (Butterworth and Punt, 1999; Punt *et al.*, 2002). A management procedure (MP; Butterworth, 2007) consists of data collection, stock status evaluation, harvest control rules (HCRs), and implementation. Fleet behaviour has generally been captured in simplistic assumptions (Kraak *et al.*, 2004), and the biological detail is often relatively plentiful (Kell and Bromley, 2004). The uncertainty about the real system and its dynamics plays a key role in evaluating management strategies. This uncertainty reflects a lack of knowledge of processes such as the stock–recruitment relationship, density-dependence, natural variability, and responses of the fleet to the measures imposed.

North Sea plaice (*Pleuronectes platessa*) and sole (*Solea solea*) are taken mainly in a mixed beam trawl fishery with a bycatch of other demersal species. Management of these two species faces many challenges because their spatial distributions and the selectivity characteristics of the gears used to catch them can have substantial side effects on the effectiveness of the measures taken in achieving the intended management objectives. For example, if the quotas set for the two species are not exhausted in synchrony, over-quota catches may be taken (and discarded or landed illegally) of the species with the most restrictive quota, which generally is depleted first. Moreover, the relative catch opportunities for the two species may differ between years. Because the mixed beam trawl fishery for flatfish presents a relatively tractable example, its problems have been well studied in the past (Rijnsdorp and Pastoors, 1995; Kell *et al.*, 1999, 2003, 2004; Pastoors *et al.*, 2000; Kell and Bromley, 2004; Kraak *et al.*, 2004).

The two target species are managed by national shares in the total allowable catch (TAC), days-at-sea restrictions on the fleet, and technical measures. The spawning-stock biomass (SSB) of plaice declined after the early 1990s and has been just above the limit biomass reference point (B_{lim}) since the mid-1990s (ICES, 2006a). The sole stock has fluctuated markedly in response to the appearance of strong year classes. The two most recent year classes are thought to be poor, which could easily reduce sole SSB below its B_{lim} in the near future (ICES, 2006c).

In 2004, the Commission of the European Communities asked the North Sea Regional Advisory Council (NSRAC) for advice "on the implementation of a recovery plan for North Sea plaice and a long-term management plan for sole". NSRAC (2005) issued its advice in July 2005, focusing only on the recovery of the plaice stock to above 230 000 t, the precautionary biomass (B_{pa}) suggested by ICES (2006c). The advice did not deal with long-term management of sole. Because the Netherlands is a major player in these flatfish fisheries, the Dutch Ministry of Agriculture, Nature Conservation and Food Quality (LNV) requested an evaluation of the likely effects of the NSRAC-proposed recovery plan on the sole stock. The details of these evaluations (Poos *et al.*, 2006) are not the focus of this manuscript; instead, we focus on the process. The main questions we pose are: (i) how has the technical evaluation of the recovery plan been initiated and carried through; (ii) how has the uncertainty in various processes been encapsulated in the OM; and (iii) how have the results been communicated and used?

Proposed NSRAC management plan

The stated objective of the management plan has been formulated for plaice only (NSRAC, 2005):

> "a multi-annual management plan should be adopted for plaice in the North Sea with an initial target of reaching an SSB at the B_{pa} level within 3–5 years with a re-evaluation after 3 years and with the long term aim of exceeding B_{pa}. The plan should be implemented as of the 1st of January 2006. The management plan is aimed at reducing pressure on juvenile plaice and would comprise structural effort reductions accompanied by stability in the TAC for plaice. The multi-annual plan should be accompanied by a monitoring and evaluation scheme, which would also include the monitoring of social and economic impact".

The basic management measure proposed was "a structural effort reduction of 15% of enforced licensed capacity limits in the international 80 mm flatfish fishery over 2006 and effort to be maintained at the new level for a further two years". There were detailed comments in the plan on exemptions for some fleets, on the way effort reductions could be embedded in national fishing plans, and on the different forms of effort reduction possible (days-at-sea regulations or decommissioning).

Despite these technical details, evaluation with the available scientific tools was difficult. For example, the plan stated: "In the event of the plaice stock falling below B_{lim} new measures would be applied". However, the nature of these measures was not described. Furthermore, the state of the sole stock was not considered, and the clients wanted to know about the effects of the plan on sole.

Such open ends appear to be a generic property of the outcome of political negotiation: specific conditions are raised, but how they should be resolved is left open until problems arise. However, for technical evaluation, open ends present difficulty in interpreting a management plan: a simulation approach requires that all actions be specified under all conditions.

When management plans are developed in conjunction with an evaluation approach, the problems can be resolved through a system of feedback loops between analysts and stakeholders. However, when an evaluation of a specific plan is requested, the open ends have to be interpreted by the analysts by formulating a range of potential scenarios.

This happened in this case: we interpreted the clause about "the new measures" in such a way that, if the perceived SSB would fall below B_{lim}, a further decrease of nominal fishing effort by 15% annually would be implemented until the perceived SSB had returned to above B_{lim}. Further, our interpretation was that the rule (although this had not been specified explicitly) would also apply to sole. An alternative scenario evaluated was without additional measures being taken when SSB would fall below B_{lim}.

The evaluation process

The parties involved in the evaluation process were the Dutch Ministry (LNV), the stakeholder flatfish Working Group (NSRAC), and the research organization (Wageningen-IMARES). The approach was to develop a FLR (www.flr-project.org; Kell *et al.*, 2007) simulation model that would allow incorporation of different hypotheses on population dynamics and the joint exploitation of the two stocks. The three parties initially interacted to define the research questions and the type of results expected, and again at the end of the process to discuss the results in the report. Between start and finish, interaction between researchers and stakeholders was negligible.

Setting up the model

The initial question in model development referred to the definition of the dimensions to be included. The dimensions reflect the type of processes included, the available information, and the dominant issues in the management plan. In this case, the partial spatial overlap between the two species suggested a two-area model (sole dominating as the target south of 56°N, plaice dominating north of 56°N), and a distinction between two main types of fleets (one each targeting sole and plaice).

The simulation model consisted of two submodels: an OM describing the biological and fleet (including economics) dynamics of the underlying system; and an MP consisting of data collection, stock assessment, and a HCR (for a graphical presentation of a similar type of simulation approach, see Kell *et al.*, 2005).

The biological processes in the OM included a stock–recruitment relationship, fixed values for natural mortality, maturity-at-age, and weight- and length-at-age, and fixed values for the relative proportion of each age group in each of the two areas. The fleet-dynamics model consisted of two beam trawl fleets: the Dutch fleet targeting primarily sole (80 mm mesh in the codend) south of 55°N, and a UK fleet targeting primarily plaice (100 mm mesh) north of 55°N. Trends in fishing effort were derived from information in the Dutch logbook database (containing records of Dutch vessels as well as UK vessels landing their catches in the Netherlands). Future effort allocations were assumed to retain the same proportions by area and fleet. Conceptually, the catching process was modelled as a combination of catchability, selectivity, effort, and technological creep (Rijnsdorp *et al.*, 2006). Catches were generated from the underlying ("true") population and split into landings and discards according to a retention ogive derived from observer trips (Van Keeken *et al.*, 2004).

The MP consisted of three main processes: sampling raw data from the underlying population; stock assessment, and short-term forecasting following standard procedures; and a HCR defining the appropriate management measure given the forecast. Sampling from the true population was mimicked by generating estimates of landings-at-age (sole) and catch-at-age (plaice; including discards), similar to the annual assessments for the two stocks. The catches were generated using the selectivity characteristics of the two fleets and a simple lognormal error with a relatively small coefficient of variation ($CV = 0.1$).

The simulation contained a "true" survey that sampled from the populations of the two species using catchability and selectivity patterns estimated from historical data in conjunction with their spatial distribution. "Observed" survey catch-at-age by species was generated by applying a lognormal error (again $CV = 0.1$), and these series were used for tuning in the stock assessment process. The stock assessment process encompassed single-species extended survivors analysis (XSA) for plaice and sole, based on catch- and landings-at-age data, respectively. XSA settings and short-term forecasts corresponded to those used by ICES (2006a).

The HCR implemented in the model attempted to mimic the NSRAC management plan. Nominal fishing effort was reduced by 15% in 2006 compared with 2005, and this level was maintained in subsequent years. In the objective stated in the plan, there is an inherent tension between reducing effort and maintaining stability of the TAC. This posed additional challenges to implementation of the model. LNV requested an additional maximum annual change in TAC of 15% to be included in the simulations, representing an extension of the stated objectives. However, the measures stated did not refer to TACs, but only to decommissioning and days-at-sea limits. Therefore, TACs would not constrain the fishery in the model anyway: the fleets simply exhausted the effort quota and reported whatever catches they generated. We did not implement a two-tier system in which either the TAC or the effort quota could constrain the fishery, because the proposed HCR did not specify how the priority between such different measures would have been set.

The annual decision process on effort quota was based on the short-term forecast of the SSB remaining after the year to which these would apply. This forecast was compared with the B_{lim} triggers defined in the plan. Implementation error with respect to misreporting or black landings was not included.

Parameter estimates

The average distribution of the species by age over the two areas was estimated from the annual beam trawl survey, which takes a synoptic sample of the plaice and sole populations in late summer (ICES, 2006b). In general, sole are mainly in the southern area, and plaice migrate from south to north with increasing age.

Information on stock trends was available for the period 1957–2004 from ICES (2006a). Recruitment estimates for the last four years were excluded because they were considered unreliable. The remaining set was used to estimate stock–recruitment relationships of the Ricker and the Beverton–Holt type. In the forward simulation, recruitment estimates were taken from the stock–recruitment relationship, taking into account the variance estimate derived from the historical relationship.

Estimating the parameters of the fleet-dynamics model was less straightforward. The interplay of simulated stock abundance and stock distributions by age and area, together with the distributions of the fleets, determines the catch profiles generated by the fleets. The empirical observations on the catch profiles of the real fleets were not sufficient to determine the parameter values for the simulation model uniquely. Therefore, any mismatch between the simulated and the real catch profiles could have been caused by incorrect assumptions in the fleet-dynamics or stock-dynamics models.

The critical question in interpreting simulation results is: how close do these have to be to reality as currently perceived? In other words, is it a problem if simulated catch profiles or stock trends differ from the historical observations? The answer depends on the type of evaluation process in which the simulation model is used. The management plan has been phrased in a stock assessment type of discourse by referring to SSB and biological reference points. Therefore, the numerical values of the variables derived from the simulations would be viewed critically against the values produced by routine stock assessment (in this case as presented by ICES, 2006a).

The combination of over-parameterization of the OM (using more parameters than there is information to estimate them from) and the requirement to generate realistic stock and fleet dynamics, implied that fixes had to be sought to generate largely comparable results. We used the technique of hindcasting to simulate the historical trajectories of stocks and fleets over the period 1995–2004. The start populations in 1995 and levels of recruitment were taken directly from ICES (2006a), and the trends in fishing effort were derived from the logbook database. Given these fixed inputs, simulations were run under different assumptions on the relative distribution of fish over the two areas, the catchabilities of the two fleets for the two species, and the increase in technical efficiency of the two fleets (Rijnsdorp *et al.*, 2006).

Some results of the exploratory analyses are shown in Figure 1; we compare the perceived values (ICES, 2006a) of landings, discards, recruitment, average fishing mortality (*F*bar), selection pattern, and SSB, the assumed relative distribution of fish over areas, and the relationship between *F*bar and effort with their simulated ("true") values, as derived from the OM. Obviously, if we wish to fix the simulated historical dynamics to match the observations, we are faced with the choice between many different options: mean selection pattern, effort–mortality relationship, landings, etc., may all be fixed, but fixing all simultaneously is impossible unless weights can be assigned to each specific parameter.

Figure 2 (left panels) shows the comparison between the (perceived as) "true" and simulated values from the basic run for four important population dynamics parameters. In the base run, all parameters were based on information from some type of analysis except for overall catchability of the fleets, which was obtained by minimizing the difference in catch-at-age between the model and as actually observed in samples from the fleet. Clearly, the simulated landings and SSB of plaice and sole are substantially smaller than those according to the assessment. Although the temporal dynamics may appear reasonably similar, a discrepancy in the overall level may have substantial effects on prospective simulations if the same absolute values are used for limit reference points. For instance, the simulated stocks will be below B_{lim} most of the time.

For plaice, an additional problem arose with the split of catches between landings and discards: simulated landings were below and simulated discards above the assessment estimates. Most plaice aged 1 and 2 years are in the southern area where the beam trawl fishery mainly operates, resulting in high simulated discards. In practice, these smaller fish may not be available to the fishery because they live in shallow coastal regions where larger vessels are not permitted. However, in a simulation environment based on two areas only, such subtle differences cannot be accounted for.

The results presented in Figure 2 (right panels) incorporate an attempt to remedy this shortcoming by manually modifying the relative spatial distribution of plaice aged 1 and 2. Because these age groups are now located more in the northern area, the overall discards decrease, so the simulated landings approach the observed landings. Also, the SSB estimates are closer. However, the simulations share a major discrepancy between simulations and assessment (Figure 1f): the linear and positive relationship

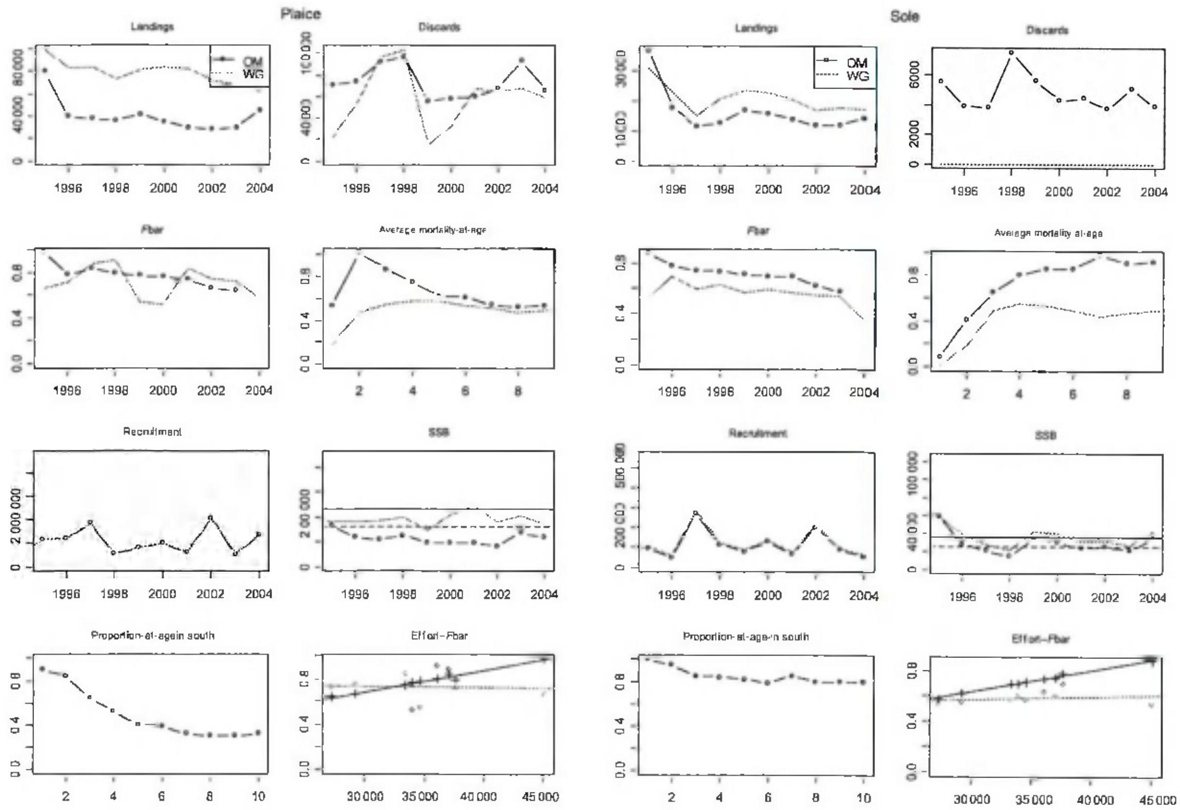

Figure 1. Comparison of hindcast population dynamics parameters for plaice (left) and sole (right) from simulations (black lines, OM) using species distribution by area from surveys and not taking into account technological creep with the corresponding estimates (red lines, Working Group) available from ICES (2006a): landings, discards, mean *F* (*F*bar), mortality-at-age, recruitment, SSB, proportion of population by age group in the southern area (input), and the relationship between effort and *F*.

between fishing effort and fishing mortality in the simulation model is a direct consequence of an assumption in the underlying model, whereas the available assessment data indicate no significant relationship at all.

Figure 3 compares the results for sole after introducing an increase in technical efficiency for the two species, as estimated by Rijnsdorp *et al.* (2006). Although the slopes of the effort–mortality relationship have become more comparable, the absolute estimates of fishing mortality in the simulation are still substantially higher than those from the assessment.

Communicating results

Because our focus during the simulations was on the assumptions behind the simulation model and the way the OM was parameterized or, more generally, on the uncertainties involved, the full range of results was presented to the parties that commissioned the study. However, those clients did not relate to the technical complexities of the modelling process. Their main frame of evaluation was whether the hindcasting process made sense (are the results similar to what has been observed?) and whether they could understand the logic of future developments (can the results be explained?). Many stakeholders read the outputs closely: if we simulated landings of plaice of 70 000 t while the TAC was just 60 000 t, they would question the validity of the results. This indicated a large degree of trust in the potential precision of the simulation models used, but at the same time a high degree of scepticism regarding their outcomes: the results were easily considered to be falsified when compared with possibly equally uncertain assessment data.

Discussion

The evaluation of the NSRAC management plan required the use of a multifleet, multistock simulation model, which is a level of complexity higher than most models used so far to evaluate fisheries management systems (Kell *et al.*, 2005, 2006b). The difficulties encountered in trying to match the results of this exercise with the type of information underlying the management advice have not been presented to highlight the modelling details, but to illustrate the general problems of such an evaluation process, of encapsulating uncertainty, and of communicating complex results.

The evaluation was carried through in an almost linear chain of actions. First, the problem was defined jointly by the clients and scientists. During the next phase, the scientists interpreted the plan, developed the model, ran the scenarios, and evaluated the results, almost in isolation. Finally, the results and conclusions were presented to the clients. The lack of interaction with the clients during the actual research phase may be explained partly by the timing of the request, because the work had to be done in December and January when the clients were heavily involved in the annual decision-making process for the following year's

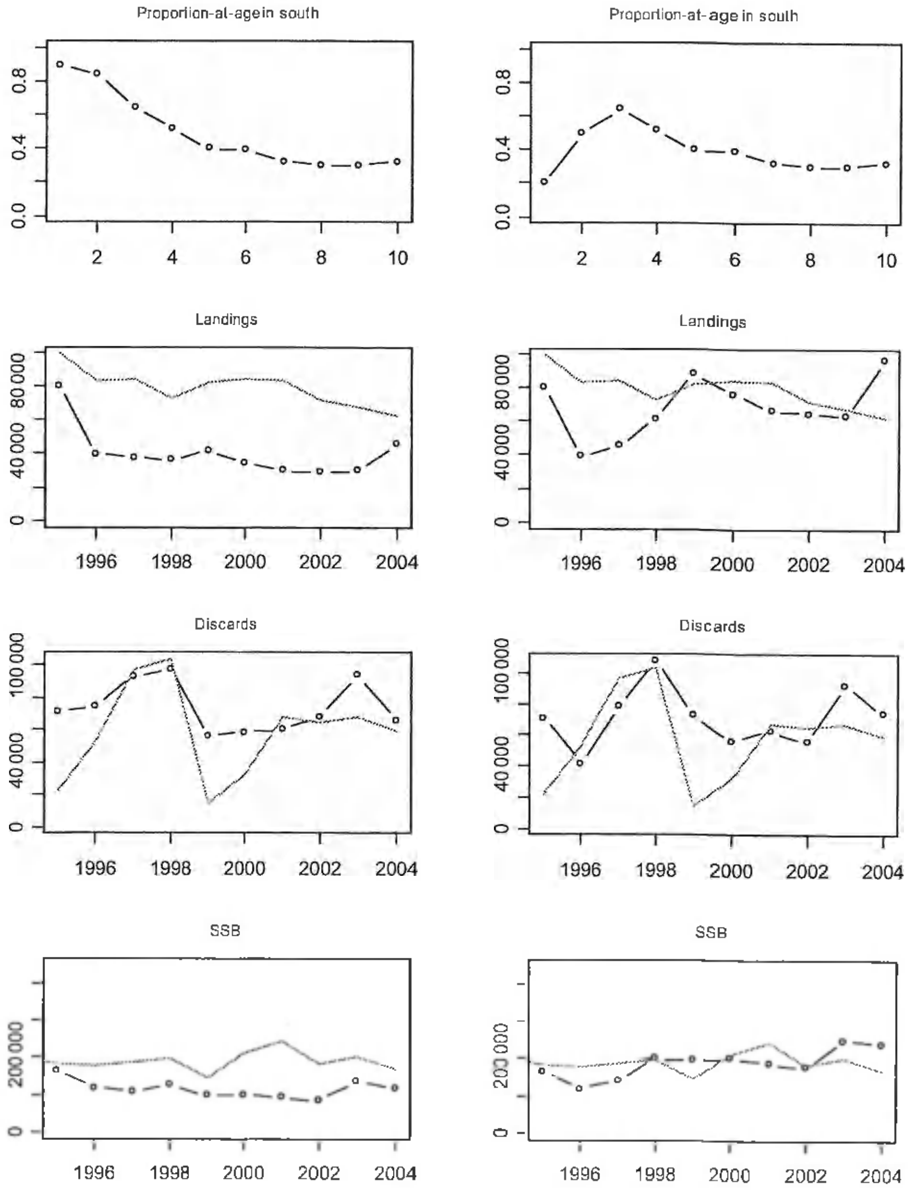

Figure 2. Comparison of hindcast landings, discards, and SSB of plaice for two patterns of the proportion-at-age present in the southern area (top panels): left panels, proportion estimated from survey data (as in Figure 1); right panels, proportion of younger ages manually modified to minimize the discrepancy in the landings.

TAC. Nevertheless, such a lack of interaction made the process less transparent (who takes responsibility for the choices made?), and no adjustments could be made to the mutual expectations of the type of results generated from the simulation model.

A generic property of management plans is that they are the product of political negotiation. Such processes often lead to an end product that describes the overall intention clearly and contains some aspects described in meticulous detail. However, other aspects are only dealt with cursorily, so requiring further interpretation. Also, specific issues may not be tractable in a simulation approach. An example is the apparent willingness of the NSRAC to consider an exception clause for the German shrimper fleet: "if scientific evidence supports the claim that plaice are not discarded in significant quantities in their targeted 80 mm sole fishery". As scientists, we did not see how this exemption could be integrated and evaluated in the simulations without having access to specific information regarding discards in that fishery.

Regarding open ends in a management plan (such as the statement that new measures might be applied under specific conditions, without actually stating what these measures might be), one must accept that the negotiation process cannot simply be reopened. However, to obtain a clearer understanding of the type of measures envisioned, it would have been beneficial for the

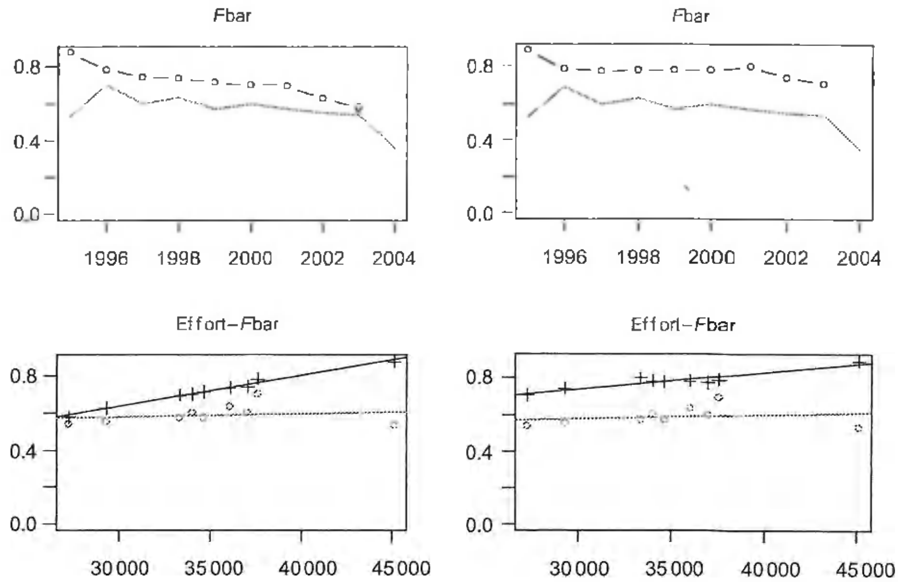

Figure 3. Comparison of hindcast average F (Fbar) and the effort–mortality relationship for sole: left panels, proportion in the south estimated by surveys and no technological creep assumed (as in Figure 1); right panels, proportion in the south manually modified (as for Figure 2) and technological creep assumed as based on Rijnsdorp *et al.* (2006).

evaluation process to have had closer and more regular interactions with the NSRAC on the class of new measures that might be explored.

The uncertainties in the OM explored in simulations addressed just three processes, but these represent a limited subset of all the processes that might influence the overall uncertainty (e.g. selectivity of the fleets, dynamic effort allocations by area, alternative relationships between selectivity, catchability, effort, and fishing mortality). This type of multidimensional model requires a strong analytical approach to explore the main factors affecting uncertainty (Kell *et al.*, 2006a). Hindcasting can be a useful tool for exploring the behavioural dynamics of a model if the dynamics can be verified by observations. As a caveat, we note that the observations available may themselves be flawed interpretations of the reality they attempt to represent. The discard observations in the North Sea plaice assessment are a good example, because the estimates are based on a modelling approach that generates huge fluctuations in discards from year to year. If simulated discards deviate from assessment discards, this may represent incorrect assumptions in either model.

Communication of results with the clients has been addressed only briefly so far. Stakeholders and fisheries managers did not appear to relate directly to the technical complexities of the modelling process, although they were concerned whether the hindcasting could reproduce the observed dynamics in the assessment underlying the TAC advice, and whether they could logically explain in their own words the developments in stocks and fisheries predicted by the simulations. We suggest that the issue of focusing too much on absolute values in the simulation output might be circumvented by expressing the results relative to an appropriately chosen reference level (for instance, the final year for which information is available). The reference points used in the management process would then also have to be scaled on that basis. Even under such a scenario, shared understanding of what the models can and cannot deliver remains an important issue. The expectations of the different parties should be laid clearly and openly on the table, to make the process as efficient as possible.

Our overall conclusion is that the simulation approach has been useful, but it might have been even more useful if used during the developmental stage of the plan, and with more interaction with the NSRAC before the plan was finally agreed.

The FLR modelling environment (Kell *et al.*, 2007) allows construction of flexible simulation models that reuse generic tools and can be attuned to different types of management rules and output measures required. However, because FLR is still in a developmental stage, the balance in the development of generic and specific tools has tended to be on the latter side. As a consequence, the resulting code has become too complex to be transferred easily to other applications. More work is needed to make the simulation models more transparent and transferable.

Acknowledgements

The study was commissioned and funded by the Dutch Ministry of Agriculture, Nature Conservation and Food Quality (LNV). The Commission of the European Communities Directorate General for Fisheries contributed to the development of the simulation model and the writing of the manuscript under the contract "Operational evaluation tools for fisheries management options" ("EFIMAS"; SSP8-CT-2003-502516).

References

Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. ICES Journal of Marine Science, 64: 613–617.

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

CEC. 2001. Proposal for a council regulation establishing measures for the recovery of cod and hake stocks. COM (2001), 724.

CEC. 2006. Communication from the Commission to the Council and the European Parliament. Implementing sustainability in EU fisheries through maximum sustainable yield. COM(2006), 360.

ICES. 2006a. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak, Copenhagen, 6–15 September 2005. ICES Document CM 2006/ACFM: 09.

ICES. 2006b. Report of the Working Group on Beam Trawl Surveys, Hamburg, 16–19 May 2006. ICES Document CM 2006/LRC: 11.

ICES. 2006c. Report of the ICES Advisory Committee on Fishery Management, Advisory Committee on the Marine Environment and Advisory Committee on Ecosystems, 2006. ICES Advice, 1–10.

Kell, L. T., and Bromley, P. J. 2004. Implications for current management advice for North Sea plaice (*Pleuronectes platessa* L.). 2. Increased biological realism in recruitment, growth, density-dependent sexual maturation and the impact of sexual dimorphism and fishery discards. Journal of Sea Research, 51: 301–312.

Kell, L. T., Cotter, A. J. R., Van Keeken, O., Pastoors, M., O'Brien, C. M., Piet, G. J., and Rackham, B. D. 2003. The influence on stock assessment advice of sampling error in research survey and international market sampling data for North Sea cod (*Gadus morhua* L.) and plaice (*Pleuronectes platessa* L.). ICES Document CM 2003/X: 5.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E. *et al.* 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640–646.

Kell, L. T., O'Brien, C. M., Smith, M. T., Stokes, T. K., and Rackham, B. D. 1999. An evaluation of management procedures for implementing a precautionary approach in the ICES context for North Sea plaice (*Pleuronectes platessa* L.). ICES Journal of Marine Science, 56: 834–845.

Kell, L. T., Pastoors, M. A., Scott, R. D., Smith, M. T., Van Beek, F. A., O'Brien, C. M., and Pilling, G. M. 2005. Evaluation of multiple management objectives for Northeast Atlantic flatfish stocks: sustainability vs. stability of yield. ICES Journal of Marine Science, 62: 1104–1117.

Kell, L. T., Pilling, G. M., Kirkwood, G. P., Pastoors, M. A., Mesnil, B., Korsbrekke, K., Abaunza, P. *et al.* 2006b. An evaluation of multi-annual management strategies for ICES roundfish stocks. ICES Journal of Marine Science, 63: 12–24.

Kell, L. T., Pilling, G. M., and Kuikka, S. 2006a. Relative importance of value of information and value of control for mixed fisheries. Poster presented at the ICES Symposium on Fisheries Management Strategies, 27–30 June 2006, Galway, Ireland.

Kell, L. T., Scott, R. D., and Hunter, E. 2004. Implications for current management advice for North Sea plaice. 1. Migration between the North Sea and English Channel. Journal of Sea Research, 51: 287–299.

Kirkwood, G. P. 1997. The revised management procedure of the International Whaling Commission. *In* Global Trends: Fisheries Management, pp. 91–99. Ed. by E. K. Pikitch, D. D Huppert, and M. P. Sissenwine. American Fisheries Society Symposium, 20.

Kraak, S. B. M., Buisman, F. C., Dickey-Collas, M., Poos, J. J., Pastoors, M. A., Smit, J. G. P., Daan, N. *et al.* 2004. How can we manage mixed fisheries? A simulation study of the effect of management choices on the sustainability and economic performance of a mixed fishery. ICES Document CM 2004/FF: 11.

McAllister, M. K., Starr, P. J., Restrepo, V. R., and Kirkwood, G. P. 1999. Formulating quantitative methods to evaluate fishery-management systems: what fishery processes should be modelled and what trade-offs should be made? ICES Journal of Marine Science, 56: 900–916.

NSRAC. 2005. Immediate Management Measures for North Sea plaice. Advice to the European Commission. http://www.nsrac.org/advices/wf20050712_Advice_on_plaice_management.pdf.

Pastoors, M. A., Rijnsdorp, A. D., and Van Beek, F. A. 2000. Effects of a partially closed area in the North Sea ("plaice box") on stock development of plaice. ICES Journal of Marine Science, 57: 1014–1022.

Poos, J. J., Machiels, M. A. M., and Pastoors, M. A. 2006. Investigation of some Management Scenarios for North Sea Sole and Plaice in 2006 and beyond. Wageningen-IMARES: CVO Report, 06.004. IJmuiden, the Netherlands.

Punt, A. E., Smith, A. D. M., and Cui, G. 2002. Evaluation of management tools for Australia's South East Fishery 1. Modelling the southeast fishery taking account of technical interactions. Marine and Freshwater Research, 53: 615–629.

Rijnsdorp, A. D., Daan, N., and Dekker, W. 2006. Partial fishing mortality per fishing trip: a useful indicator of effective fishing effort in mixed demersal fisheries. ICES Journal of Marine Science, 63: 556–566.

Rijnsdorp, A. D., and Pastoors, M. A. 1995. Modelling the spatial dynamics and fisheries of North Sea plaice (*Pleuronectes platessa* L.) based on tagging data. ICES Journal of Marine Science, 52: 963–980.

Van Keeken, O. A., Quirijns, F. J., and Pastoors, M. A. 2004. Analysis of Discarding in the Dutch Beam Trawl Fleet. RIVO Report, C034/04. IJmuiden, the Netherlands.

doi:10.1093/icesjms/fsm051